

0590  
1009

OIPE

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/940,673

DATE: 09/26/2001  
TIME: 12:31:54

Input Set : N:\Crf3\RULE60\09940673.txt  
Output Set: N:\CRF3\09262001\I940673.raw

#2

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Gorski, David H.  
6 Walsh, Kenneth

8 (ii) TITLE OF INVENTION: Growth Arrest Homeobox Gene

10 (iii) NUMBER OF SEQUENCES: 19

12 (iv) CORRESPONDENCE ADDRESS:

13 (A) ADDRESSEE: Calfee, Halter, and Griswold

14 (B) STREET: 800 Superior Avenue

15 (C) CITY: Cleveland

16 (D) STATE: Ohio

17 (E) COUNTRY: U.S.A.

18 (F) ZIP: 44114-2688

20 (v) COMPUTER READABLE FORM:

21 (A) MEDIUM TYPE: Floppy disk

22 (B) COMPUTER: IBM PC compatible

23 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

24 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25

26 (vi) CURRENT APPLICATION DATA:

27 (A) APPLICATION NUMBER: US/09/940,673

C--> 28 (B) FILING DATE: 27-Aug-2001

29 (C) CLASSIFICATION:

31 (vii) PRIOR APPLICATION DATA:

32 (A) APPLICATION NUMBER: 09/078,465

33 (B) FILING DATE: 23-SEPT-1997

35 (viii) ATTORNEY/AGENT INFORMATION:

36 (A) NAME: Golrick, Mary E.

37 (B) REGISTRATION NUMBER: 34829

38 (C) REFERENCE/DOCKET NUMBER: 22311/00114

40 (ix) TELECOMMUNICATION INFORMATION:

41 (A) TELEPHONE: (216) 622-8200

42 (B) TELEFAX: (216) 241-0816

43 (C) TELEX: 980499

46 (2) INFORMATION FOR SEQ ID NO: 1:

48 (i) SEQUENCE CHARACTERISTICS:

49 (A) LENGTH: 2244 base pairs

50 (B) TYPE: nucleic acid

51 (C) STRANDEDNESS: both

52 (D) TOPOLOGY: linear

54 (ii) MOLECULE TYPE: cDNA

56 (iii) HYPOTHETICAL: NO

58 (iv) ANTI-SENSE: NO

61 (ix) FEATURE:

62 (A) NAME/KEY: CDS

63 (B) LOCATION: 197..1108

66 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

68 GTCAAAGTGT TATACTGCA GGAGACTGGC CGCTCGGCTC AGGACTGGGA TTAGCGGGCT

ENTERED

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/940,673

DATE: 09/26/2001  
TIME: 12:31:54

Input Set : N:\Crf3\RULE60\09940673.txt  
Output Set: N:\CRF3\09262001\I940673.raw

70	CTGCTCAAAC	CCGCGCGGCT	TTTACATTAG	GAGTGAGTGG	GGGAGAGTCC	TAGGATTCT	120
72	AGTGAAAGT	GACAGCGCTT	GGTGGACTTT	GGGACCTTCG	TGAAGTCTTC	TGCTTGGAAAG	180
74	CTGAGACTTG	CATGCC	ATG GAA CAC CCC	CTC TTT	GGC TGC CTG CGC	AGC	229
75	Met Glu His Pro Leu Phe Gly Cys Leu Arg Ser						
76	1	5	10				
78	CCC CAC GCC ACA GCG CAA GGC TTG CAC CCC TTC TCG CAG TCT TCT CTG						277
79	Pro His Ala Thr Ala Gln Gly Leu His Pro Phe Ser Gln Ser Ser Leu						
80	15	20	25				
82	GCC CTC CAT GGA AGA TCT GAC CAC ATG TCC TAC CCC GAA CTC TCC ACA						325
83	Ala Leu His Gly Arg Ser Asp His Met Ser Tyr Pro Glu Leu Ser Thr						
84	30	35	40				
86	TCT TCC TCG TCT TGC ATA ATC GCG GGA TAC CCC AAT GAG GAG GGC ATG						373
87	Ser Ser Ser Cys Ile Ile Ala Gly Tyr Pro Asn Glu Glu Gly Met						
88	45	50	55				
90	TTT GCC AGC CAG CAT CAC AGG GGG CAC CAC CAC CAC CAC CAC CAC CAC						421
91	Phe Ala Ser Gln His His Arg Gly His His His His His His His His						
92	60	65	70	75			
94	CAT CAC CAC CAC CAG CAG CAG CAG CAC CAG GCT CTG CAA AGC AAC TGG						469
95	His His His Gln Gln Gln His Gln Ala Leu Gln Ser Asn Trp						
96	80	85	90				
98	CAC CTC CCG CAG ATG TCC TCC CCG CCA AGC GCG GCC CGG CAC AGC CTT						517
99	His Leu Pro Gln Met Ser Ser Pro Pro Ser Ala Ala Arg His Ser Leu						
100	95	100	105				
102	TGC CTG CAG CCT GAT TCC GGA GGG CCC CCG GAG CTG GGG AGC AGC AGC CCT						565
103	Cys Leu Gln Pro Asp Ser Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro						
104	110	115	120				
106	CCG GTC CTG TGC TCC AAC TCT TCT AGC CTG GGC TCC AGC ACC CCG ACC						613
107	Pro Val Leu Cys Ser Asn Ser Ser Leu Gly Ser Ser Thr Pro Thr						
108	125	130	135				
110	GGA GCC GCG TGC GCA CCA AGG GAT TAT GGC CGT CAA GCG CTG TCA CCC						661
111	Gly Ala Ala Cys Ala Pro Arg Asp Tyr Gly Arg Gln Ala Leu Ser Pro						
112	140	145	150	155			
114	GCA GAA GTG GAG AAG AGA AGT GGC AGC AAA AGA AAA AGC GAC AGT TCA						709
115	Ala Glu Val Glu Lys Arg Ser Gly Ser Lys Arg Lys Ser Asp Ser Ser						
116	160	165	170				
118	GAT TCC CAG GAA AAT TAC AAG TCA GAA GTG AAC AGC AGC AAA CCT AGG						757
119	Asp Ser Gln Glu Gly Asn Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg						
120	175	180	185				
122	AGG GAA AGA ACA GCT TTC ACC AAA GAG CAA ATC AGA GAA CTT GAG GCA						805
123	Arg Glu Arg Thr Ala Phe Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala						
124	190	195	200				
126	GAG TTC GCC CAT CAT AAC TAT CTG ACC AGA CTG AGA AGA TAT GAG ATA						853
127	Glu Phe Ala His His Asn Tyr Leu Thr Arg Leu Arg Arg Tyr Glu Ile						
128	205	210	215				
130	GCG GTG AAC CTA GAC CTC ACT GAA AGA CAG GTG AAA GTG TGG TTC CAG						901
131	Ala Val Asn Leu Asp Leu Thr Glu Arg Gln Val Lys Val Trp Phe Gln						
132	220	225	230	235			
134	AAC AGG AGA ATG AAG TGG AAG CGG GTC AAG GGG GGA CAA CAA GGA GCT						949
135	Asn Arg Arg Met Lys Trp Lys Arg Val Lys Gly Gly Gln Gly Ala						

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/940,673

DATE: 09/26/2001  
TIME: 12:31:54

Input Set : N:\Crf3\RULE60\09940673.txt  
Output Set: N:\CRF3\09262001\I940673.raw

136	240	245	250														
138	GCA	GCC	CGA	997													
139	GAA	AAG	GAA														
140	CTG	GTG	AAT														
141	AAA	AAG	GGA														
142	ACA	CTT	CTT														
143	Leu	Leu	Val														
144	Asn	Val	Lys														
145	Lys	Gly	Thr														
146	Leu	Leu	Leu														
147	Ala	Arg	Arg														
148	255	260	265														
149	270	275	280														
150	285	290	295														
151	CAC	GCA	CAC	1093													
152	TTA	TGATA	CATAC														
153	AGAGACCAGC	TCCGTTCTCA	GGAAAGCACC														
154	His	Ala	His	1145													
155	Leu																
156	300																
157	ATTGTGATGG	CAAATCTCAC	CCAAACATCG	1205													
158	TTAACACG	GCATCTCAAG	TTCTGTTCTC	1265													
159	GGAGGTTAC	ATGATTGATA	GAAGGTTAC														
160	ACTAAGTGCC	TCTTATTGAA	GATGCTTCCA	1325													
161	AAACATATCTA	CAGTCAAATT	GGAGAAAGTG														
162	160	161	162														
163	AAATTAATT	GCTACCAAGA	GCACACTCGG	1385													
164	TAAGACATT	TGACTCAAGT	TGTCTCCAGA	1445													
165	164	165	166														
166	GTGAAGATGT	TATAGAAATG	CTTCAGTTGTA	1505													
167	CTTGAACAT	CCAGGTCAATG	TGTGTGACAC														
168	166	167	168														
169	TGGGCAGGTA	TTTGCTTTG	CTTGCACAGA	1565													
170	AACTAAACT	GACTATCAAGT	TAACCCATGA														
171	168	169	170														
172	ATAGTTTAT	CTTGAAACAGC	GGATGCTCT	1625													
173	TTTGCACAGC	GGATGCTCT	GGATGGATATA														
174	AAATGAAACTG	AAATGAACTG	AAATGAACTG	1685													
175	174	176	177														
176	AAATTCTGTA	TATATTACTC	CTAACATTT	1745													
177	CTAACATTT	TTCCCTGTCTT	TTCCCTGTCTT														
178	176	177	178														
179	AAATTCTGTA	TTCCCTGTCTT	TTCCCTGTCTT	1805													
180	TTCCCTGTCTT	TTCCCTGTCTT	TTCCCTGTCTT														
181	ATCGCTCGAA	GACATTAGCG	TCATAGGCAA	1865													
182	180	181	182														
183	GCAGATGTCT	GATTTGTAT	CTTTAAACTG	1925													
184	182	183	184														
185	TTGTTTATGA	TTGTTTATGA	TTGTTTATGA	1985													
186	184	185	186														
187	AAAGCTAGT	TCTTCAGAG	TTGTTTATGA	2045													
188	186	187	188														
189	AAATACATTA	TTGTTTATGA	TTGTTTATGA	2105													
190	188	189	190														
191	AAAAAAAGTT	AAATAAATG	AAATAAATG	2165													
192	190	191	192														
193	(2)	INFORMATION FOR SEQ ID NO: 2:		2225													
194	(i)	SEQUENCE CHARACTERISTICS:															
195	(A)	LENGTH: 303 amino acids															
196	(B)	TYPE: amino acid															
197	(D)	TOPOLOGY: linear															
198	(ii)	MOLECULE TYPE: protein															
199	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: 2:															
200	Met	Glu	His	Pro	Leu	Phe	Gly	Cys	Leu	Arg	Ser	Pro	His	Ala	Thr	Ala	
201	1	5	10	15													
202	Gln	Gly	Leu	His	Pro	Phe	Ser	Gln	Ser	Ser	Leu	Ala	Leu	His	Gly	Arg	
203	20	25	30														
204	Ser	Asp	His	Met	Ser	Tyr	Pro	Glu	Leu	Ser	Thr	Ser	Ser	Ser	Ser	Cys	
205	35	40	45														
206	Ile	Ile	Ala	Gly	Tyr	Pro	Asn	Glu	Glu	Gly	Met	Phe	Ala	Ser	Gln	His	
207	50	55	60														
208	65	70	75	80													

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/940,673

DATE: 09/26/2001  
TIME: 12:31:54

Input Set : N:\Crf3\RULE60\09940673.txt  
Output Set: N:\CRF3\09262001\I940673.raw

219 Gln Gln Gln His Gln Ala Leu Gln Ser Asn Trp His Leu Pro Gln Met  
220 85 90 95  
222 Ser Ser Pro Pro Ser Ala Ala Arg His Ser Leu Cys Leu Gln Pro Asp  
223 100 105 110  
225 Ser Gly Gly Pro Pro Glu Leu Gly Ser Ser Pro Pro Val Leu Cys Ser  
226 115 120 125  
228 Asn Ser Ser Ser Leu Gly Ser Ser Thr Pro Thr Gly Ala Ala Cys Ala  
229 130 135 140  
231 Pro Arg Asp Tyr Gly Arg Gln Ala Leu Ser Pro Ala Glu Val Glu Lys  
232 145 150 155 160  
234 Arg Ser Gly Ser Lys Arg Lys Ser Asp Ser Ser Asp Ser Gln Glu Gly  
235 165 170 175  
237 Asn Tyr Lys Ser Glu Val Asn Ser Lys Pro Arg Arg Glu Arg Thr Ala  
238 180 185 190  
240 Phe Thr Lys Glu Gln Ile Arg Glu Leu Glu Ala Glu Phe Ala His His  
241 195 200 205  
243 Asn Tyr Leu Thr Arg Leu Arg Arg Tyr Glu Ile Ala Val Asn Leu Asp  
244 210 215 220  
246 Leu Thr Glu Arg Gln Val Lys Val Trp Phe Gln Asn Arg Arg Met Lys  
247 225 230 235 240  
249 Trp Lys Arg Val Lys Gly Gly Gln Gln Gly Ala Ala Ala Arg Glu Lys  
250 245 250 255  
252 Glu Leu Val Asn Val Lys Lys Gly Thr Leu Leu Pro Ser Glu Leu Ser  
253 260 265 270  
255 Gly Ile Gly Ala Ala Thr Leu Gln Gln Thr Gly Asp Ser Leu Ala Asn  
256 275 280 285  
258 Asp Asp Ser Arg Asp Ser Asp His Ser Ser Glu His Ala His Leu  
259 290 295 300

262 (2) INFORMATION FOR SEQ ID NO: 3:

264 (i) SEQUENCE CHARACTERISTICS:  
265 (A) LENGTH: 941 base pairs  
266 (B) TYPE: nucleic acid  
267 (C) STRANDEDNESS: both  
268 (D) TOPOLOGY: linear  
270 (ii) MOLECULE TYPE: cDNA  
272 (iii) HYPOTHETICAL: NO  
274 (iv) ANTI-SENSE: NO  
277 (ix) FEATURE:  
278 (A) NAME/KEY: CDS  
279 (B) LOCATION: 33..941

282 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

284 GTCTTCTACC TGGAACCCGA AACTTGCATG CT ATG GAA CAC CCG CTC TTT GGC	53
285 Met Glu His Pro Leu Phe Gly	
286 1 5	
288 TGC CTG CGC AGC CCT CAC GCC ACG GCG CAA GGC TTG CAC CCG TTC TCC	101
289 Cys Leu Arg Ser Pro His Ala Thr Ala Gln Gly Leu His Pro Phe Ser	
290 10 15 20	
292 CAA TCC TCT CTC GCC CTC CAT GGA AGA TCT GAC CAT ATG TCT TAC CCC	149
293 Gln Ser Ser Leu Ala Leu His Gly Arg Ser Asp His Met Ser Tyr Pro	

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/940,673

DATE: 09/26/2001  
TIME: 12:31:54

Input Set : N:\Crf3\RULE60\09940673.txt  
Output Set: N:\CRF3\09262001\I940673.raw

294	25	30	35	
296	GAG CTC TCT ACT TCT TCC TCA TCT TGC ATA ATC GCG GGA TAC CCC AAC			197
297	Glu Leu Ser Thr Ser Ser Ser Cys Ile Ile Ala Gly Tyr Pro Asn			
298	40	45	50	55
300	GAA GAG GAC ATG TTT GCC AGC CAG CAT CAC AGG GGG CAC CAC CAC CAC			245
301	Glu Glu Asp Met Phe Ala Ser Gln His His Arg Gly His His His His			
302	60	65	70	
304	CAC CAC CAC CAT CAC CAC CAT CAG CAG CAG CAC CAG GCT CTG CAA			293
305	His His His His His Gln Gln Gln Gln His Gln Ala Leu Gln			
306	75	80	85	
308	ACC AAC TGG CAC CTC CCG CAG ATG TCT TCC CCA CCG AGT GCG GCT CGG			341
309	Thr Asn Trp His Leu Pro Gln Met Ser Ser Pro Pro Ser Ala Ala Arg			
310	90	95	100	
312	CAT AGC CTC TGC CTC CAG CCC GAC TCT GGA GGG CCC CCA GAG TTG GGG			389
313	His Ser Leu Cys Leu Gln Pro Asp Ser Gly Gly Pro Pro Glu Leu Gly			
314	105	110	115	
316	AGC AGC CCG CCC GTC CTG TGC TCC AAC TCT TCC AGC TTG GGC TCC AGC			437
317	Ser Ser Pro Pro Val Leu Cys Ser Asn Ser Ser Leu Gly Ser Ser			
318	120	125	130	135
320	ACC CCG ACT GGG GCC GCG TGC GCG CCG GGG GAC TAC GGC CGC CAG GCA			485
321	Thr Pro Thr Gly Ala Ala Cys Ala Pro Gly Asp Tyr Gly Arg Gln Ala			
322	140	145	150	
324	CTG TCA CCT GCG GAG GCG GAG AAG CGA AGC GGC GGC AAG AGG AAA AGC			533
325	Leu Ser Pro Ala Glu Ala Glu Lys Arg Ser Gly Gly Lys Arg Lys Ser			
326	155	160	165	
328	GAC AGC TCA GAC TCC CAG GAA GGA AAT TAC AAG TCA GAA GTC AAC AGC			581
329	Asp Ser Ser Asp Ser Gln Glu Gly Asn Tyr Lys Ser Glu Val Asn Ser			
330	170	175	180	
332	AAA CCC AGG AAA GAA AGG ACA GCA TTT ACC AAA GAG CAA ATC AGA GAA			629
333	Lys Pro Arg Lys Glu Arg Thr Ala Phe Thr Lys Glu Gln Ile Arg Glu			
334	185	190	195	
336	CTT GAA GCA GAA TTT GCC CAT CAT AAT TAT CTC ACC AGA CTG AGG CGA			677
337	Leu Glu Ala Glu Phe Ala His His Asn Tyr Leu Thr Arg Leu Arg Arg			
338	200	205	210	215
340	TAC GAG ATA GCA GTG AAT CTG GAT CTC ACT GAA AGA CAG GTA AAA GTC			725
341	Tyr Glu Ile Ala Val Asn Leu Asp Leu Thr Glu Arg Gln Val Lys Val			
342	220	225	230	
344	TGG TTC CAA AAC AGG CGG ATG AAG TGG AAG AGG GTA AAG GGT GGA CAG			773
345	Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Arg Val Lys Gly Gly Gln			
346	235	240	245	
348	CAA GGA GCT GCG GCT CGG GAA AAG GAA CTG GTG AAT GTG AAA AAG GGA			821
349	Gln Gly Ala Ala Ala Arg Glu Lys Glu Leu Val Asn Val Lys Lys Gly			
350	250	255	260	
352	ACA CTT CTC CCA TCA GAG CTG TCG GGA ATT GGT GCA GCC ACC CTC CAG			869
353	Thr Leu Leu Pro Ser Glu Leu Ser Gly Ile Gly Ala Ala Thr Leu Gln			
354	265	270	275	
356	CAA ACA GGG GAC TCT ATA GCA AAT GAA GAC AGT CAC GAC AGT GAC CAC			917
357	Gln Thr Gly Asp Ser Ile Ala Asn Glu Asp Ser His Asp Ser Asp His			
358	280	285	290	295

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/940,673

DATE: 09/26/2001

TIME: 12:31:55

Input Set : N:\Crf3\RULE60\09940673.txt

Output Set: N:\CRF3\09262001\I940673.raw

L:27 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:28 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]